Online Supplemental Materials

for

To Punish or Repair?

Evolutionary Psychology and Lay Intuitions about Modern Criminal Justice

Additional Methods Descriptions – Study 1

Subjects and procedures

The subjects were 4,116 adolescent Danes completing a general or vocational upper-secondary education. The paper-and-pencil questionnaire was filled out during class time. 2076 subjects were males and 2030 were females (10 subjects did not report their sex). The mean age was 18.25 (with a standard deviation of 2.79). The schools were situated in different geographical locations, had different ethnic compositions, degrees of urbanization, and students from a wide range of different social backgrounds (approximately 85 percent of all Danish adolescents receive an upper secondary education). The full wordings of the vignettes are presented in Table SI1.

-TABLE SI1 ABOUT HERE -

Measures

All variables are coded between 0 and 1. High values indicate high perceived association value, high perceived seriousness, a preference for an intense sanction and support for reparative over punitive sanctions, respectively.

Perceived association value. Perception of association value was operationalized by an item describing claims by two individuals, A and B. The claims were as follows: A says 'I believe this man can realize his mistake, so that he won't do it again.' B says: 'When a man does something like

that, he will surely do it again'. Subjects were asked to state their opinion on a five-point scale ranging from "fully agree with A" to "fully agree with B".

Perceived seriousness of the crime. In line with previous research (cf. Stylianou, 2004), the perceived seriousness of the crime was measured by asking the subjects to indicate this on an eleven-point scale with the poles defined as 'not very serious' and 'very serious'.

Preferred intensity of sanction measure. As is common in the literature, the preferred intensity of the sanction was measured by asking the respondents to choose among a list of specific reactions of increasing severity, from a warning to prison time. Specifically, the respondents could choose between a warning, a fine/compensation, three community service sentences with increasing durations and three prison sentences with increasing durations. The durations were based on the average prison sentences for the specific type of crime in Denmark (as reported by the Danish Ministry of Justice, 2003). The answers form an eight-point scale ranging from low to high intensity of the reaction.

Preference for punitive or reparative sanction. Whether a punitive or reparative sanction is preferred was measured by an item involving two claims and a five-point scale ranging from "fully agree with A" to "fully agree with B". The claims were: A says: 'The important thing is that this man is being punished for his deeds'. B says: 'It is more important that this man is being helped to realize that he has caused harm'. These claims cover a punitive and a reparative reaction, respectively. The measure is coded such that a high value indicates support for reparative sanctions over punitive sanctions.

Outgroup hostility measure. Finally, general outgroup hostility is a summation of answers to two items based on a five-point Likert-scale answer-format ($\alpha = 0.83$). The two items were "Current immigration poses a serious threat to Danish culture" and "We should not allow any more immigrants into Denmark". A high value indicates a high level of hostility.

Additional Methods Descriptions – Study 2

Subjects and procedures

138 undergraduates at a large American research university completed a 20 minutes survey in exchange for course credits. 68 subjects were males and 69 were females (1 subject did not report sex). The mean age was 18.6 (with a standard deviation of .85). The full wording of the experimental vignette is presented in Table SI2.

-TABLE SI1 ABOUT HERE -

Measures

High values indicate high perceived association value, high perceived seriousness, a preference for an intense sanction and support for reparative over punitive sanctions, respectively.

       Perceived association value: productivity Our primary measure of perception of association value was operationalized by the item: "In your opinion, how likely is it that this criminal can someday become a productive member of society?". Respondents were asked to indicate their level of agreement on a five-point scale ranging from 'Extremely likely' to 'Not likely at all'.

       Perceived association value: recidivism. Our second measure of association value in Study 2, perceptions of likelihood of recidivism, was operationalized by the item: "In your opinion, how likely is it that this criminal, when released from jail, will commit a crime again?" Respondents were asked to indicate their level of agreement on a five-point scale ranging from 'Extremely likely' to 'Not likely at all'.

       Perceived seriousness of the crime. The perceived seriousness of the crime was measured by asking the subjects the following question: "In your opinion, how serious was the

crime you read about?". Answers were recorded on a five-point scale from 'Extremely serious' to 'Not serious at all'.

Preferred intensity of sanction measure. The preferred intensity of the sanction was measured by asking the respondents: "How severe a punishment should be given for this offense?". Again, answers were recorded on a five-point scale ranging from 'Extremely severe' to 'Not severe'.

Preference for punitive or reparative sanction. To provide a more demanding measure of whether a punitive or reparative sanction is preferred, we asked two questions about reparative sanctions: "Rehabilitation programs – such as job training or college degree programs – are often available to convicted criminals. Would you recommend that this criminal participate in such a rehabilitation program?" and "Drug and alcohol treatment programs are often available to convicted criminals. If this criminal was found to have a drug or alcohol problem, would you recommend that he take part in a drug treatment program?". Subjects could choose between recommending these programs instead of prison, in addition to prison or not at all. Based on answers to these questions, we created a dichotomous measure of punitive versus reparative preference. All subjects that indicated that they preferred that the criminal take part in a rehabilitation or treatment program *instead* of serving a prison sentence were coded as having a reparative preference (a value of 1). All others were coded as having a punitive preference (a value of 0).

Data analysis

Some of the analyses require the use of linear regression models to test mediation and interaction effects. Therefore, to make the effect sizes comparable across analyses our studies, we – whenever appropriate – use linear regression to test our predictions. In linear regression, the superior measure of effect sizes is simply the unstandardized regression coefficients (b) computed from variables

coded between 0 and 1 (Achen, 1982: 76-77). Hence, in all cases, variables are coded between 0 and 1. These effect sizes vary between -1 and 1 and have an interpretation as the change in percentages of the full scale of variable Y when variable X changes from its minimum to its maximum value. For example, an effect size of .2 signifies that a change in X from its minimum (0) to its maximum (1) value cause a change in Y of 20 percent of the full scale (e.g., from .35 to .55).

Additional Analyses

Study 1: Vignette 1

In the main text reporting of vignette 1, we pool data from two experimental conditions. Hence, vignette 1 included an experimental manipulation that simultaneously manipulated cues relevant to the computation of future association value (e.g., expression of remorse in condition 2) and cues relevant to computation of current welfare-trade ratios (e.g., increased cost-imposition in condition 1). According to the recalibrational theory, this manipulation should both affect sanctioning severity and preference for a punitive versus a reparative response. Consistent with this, we find that the experimental manipulation both affects the preferred intensity of the sanction (partial r=.29; p < .001; controlling for preference for a reparative versus punitive sanction) such that more intense sanctions are preferred in the face of increased cost-imposition (condition 1). Furthermore, the experimental manipulation also influences the preference for a reparative over a punitive reaction (partial r=.26, p < .001; controlling for preferred intensity of the sanction).

Study 1: Harm and seriousness

In our analyses, we have focused on causes and effects of association values, which constitute an augmentation of our ability to predict and explain criminal justice intuitions. As described above, a large criminological literature has documented that crime seriousness is based on the harm inflicted.

As a final analysis, we can compare the means of our three vignettes to see whether they replicate this well-established pattern. On a 0-1 scale, the vandalism case had a mean seriousness rating of 0.39 (s.d. = 0.23), the robbery with assault had a mean rating of 0.74 (s.d. = 0.18), and the rape was rated 0.91 (s.d. = 0.20). All differences are significant at the 0.001-level. Hence, this pattern conforms to results in the extant literature, where most people view vandalism, assault, and rape as inflicting increasing levels of harm (e.g., Rossi et al., 1974). Association values were computed from a diverse range of cues, but these had little or no effect on judgments of the crime's seriousness. Computations of crime seriousness appear to more narrowly focus on harm.

We note that this is consistent with results by Lieberman and Linke (2007), showing that judgments of moral wrongness—which (presumably) track how serious the crime is seen to be—do not vary depending on whether the criminal is a family member, ingroup member (schoolmate), or outgroup member (foreigner).

In contrast, these cues to association value did have an effect on the intensity of the sanction preferred for the non-violent property crimes that Lieberman and Linke used. This is consistent with our results, especially from vignette 2 (vandalism with no violence). Although we found that intensity of sanction was primarily determined by the crime's seriousness, there was a (smaller) effect of association value on intensity.

Lastly, Lieberman and Linke's results suggest that people with higher association value are seen as more easily recalibrated: their subjects expected family members to feel more remorse the crime than schoolmates, and schoolmates more than foreigners. Profession of remorse is one of three linked cues to association value manipulated in Study 2.

Study 2: Robustness of meditational analysis using SEM

Study 2 is based on a much smaller sample than Study 1. Given this, p-values are higher in Study 2 even if effects sizes are similar to those reported in Study 1. These features of Study 2 make it relevant to demonstrate the robustness of the basic results using other statistical techniques. In particular, given that the study analyses the effects of indirect paths, it is relevant to use other techniques designed for meditational analyses. Here, we show that the results are robust to the use of two other statistical packages for meditational analyses available in the statistical software, Stata.

While the use of the command, binary_mediation, in Stata (used in the manuscript text) is recommended when analyzing indirect effects on a binary outcome variable (Stata FAQ, 2012), there is also another command, medeff, available (Hicks & Tingley, 2011). Using medeff provides substantially similar and statistical significant effects. Specifically, we have tested whether we are able to replicate that perceptions of productivity mediate the effect of the experimental manipulation on preferences for reparative over punitive reactions. All values of the 90 per cent confidence interval for the mediation effect (i.e., corresponding to the one-tailed test for the existence of a mediation effect) are positive [.002; .063], indicating a significant effect at the .05-level (one-tailed).

Another common technique for analyzing indirect paths is Structural Equation Modeling (SEM). A key challenge here, however, is that in the SEM module available in Stata does not to allow the researcher to utilize techniques tailored for binary dependent variables (e.g., logitistic models). Hence, when using SEM we have been required to treat the binary variable as continuous. It is debated how large the problems associated with this violation are (for an optimistic view, see Hellevik, 2009). Either way, we have build a structural model based on the main paths from Figure 4 in the main text. The model together with its path coefficients is shown in Figure SI1. As can be seen, we can replicate all direct paths with an adequate fit. Furthermore, using SEM, we also find a

significant indirect path from the experimental manipulation through perceptions of association value on preferences for sanctioning type (coeff. = .04, p = .04, one-tailed).

Together these analyses show that the findings reported in the main text are robust to the use of other analytical and statistical techniques.

Study 2: Measures of association value

In Study 1 perceived association value of the criminal was gauged by asking about the criminal's perceived likelihood of recidivism. A primary goal of Study 2 was to measure perceived association value by focusing on the perceived future productivity of the criminal. However, to assess the relationship between these two constructs, we also included a measure of perceived recidivism in Study 2.

We find that the two measures of association value are highly correlated (r =.50, p<.001). This seems to indicate that they are tracking closely related psychological constructs. Furthermore, we can replicate the findings reported in the main text in study 2, using a scale that combines these two measures of association values. Hence, the combined measure is significantly mediating the effect of the experimental manipulation of association value on preference for reparative over punitive sanctions. The coefficient of the indirect path is .09 and the bias corrected 90 % two-tailed confidence interval (corresponding to .05-level with one-tailed tests) is [.03; .24].

A recent study has suggested that productivity (relationship value) and recidivism (exploitation risk) are psychologically represented as two separate things (Burnette et al., 2012). Our analysis, in contrast, suggests that they are closely related. For the present theory, however, the most important thing is that both constructs are expected to track the net future value of associating with an individual and, hence, that both constructs influence decisions about whether to punish or reconcile with exploiters. This is the case in both the present study and in the study by Burnette et

al. (2012). In this way, computed association value is perhaps best viewed as downstream summary

variable of other more basic computations including expected benefits (productivity) and expected

costs (recidivism) (see Petersen et al., 2010).

Supplemental References

Achen, Christopher H. (1982). *Interpreting and using regression.* Sage university papers: Quantitative applications in the social sciences, no. 07-029. Newbury Park: Sage Publications, Inc.

Burnette, Jeni L., Michael E. McCullough, Daryl R. Van Tongeren, &Don E. Davis (2012). " Forgiveness Results From Integrating Information About Relationship Value and Exploitation Risk". *Personality and Social Psychology Bulletin*, 38: 345-356.

Danish Ministry of Justice (2003). *Strafniveau, strafskærpelser og kapacitet*, København: Justitsministeriets Forskningsenhed.

Hellevik, Ottar (2009). "Linear versus logistic regression when the dependent variable is a dichotomy". *Quality and Quantity*, 43 (1): 59-74.

Hicks, Raymond, & Dusin Tingley (2011). "Causal Mediation Analysis". *The Stata Journal*, 11 (4): 1-15.

Stata FAQ (2012). "How can I perform mediation with binary variables?". UCLA Academic Technology Services, http://www.ats.ucla.edu/stat/stata/faq/binary_mediation.htm, accessed April 22 2012.

Stylianou, S. (2003). "Measuring crime seriousness perceptions: What have we learned and what else do we want to know", *Journal of Criminal Justice,* 31, 37-56.

Table SI1. Descriptions of vignettes and experimental manipulations in Study 1.

| Vignette 1	
Condition 1	Condition 2
A man has been arrested for assaulting a young woman on her way home from work. The assault took place in broad daylight and the man forced the young woman to hand over her purse and jewelry. Following the assault, the young woman was off work because of psychological after-effects.	A young man has been arrested for assault. The assault took place in broad daylight and the man forced the victim to hand over her purse and jewelry. During the police's interrogations the young man suffer a break down and tells that he needed the money to pay a local gang that were blackmailing him.
Vignette 2	
Condition 1	Condition 2
A man has been arrested for breaking a shop window during a night out. He was very drunk and could not really remember anything. He has been found guilty of similar offences three times before.	A man has been arrested for breaking a shop window during a night on the town. He was very drunk and could not really remember anything. He has no criminal record.
Vignette 3	
Condition 1	Condition 2
A young man doing his compulsory military service has been sentenced for rape. The female victim had received a drink from him at a night club, and afterwards he accompanied her home. The woman described how the draftee forced his way into her apartment and raped her.	A young man with immigrant background has been sentenced for rape. The female victim had received a drink from him at a night club, and afterwards he accompanied her home. The woman described how the immigrant forced his way into her apartment and raped her.

Notes. All subjects were presented with all three vignettes but, for each vignette, they were randomly assigned to one of

the conditions. Analysis of vignette 1 reported in the article are based on pooled data from conditions 1 and 2. See

additional analyses in this supplemental information for further discussion.

Table SI2. Description of vignette and experimental manipulation in Study 2.

Conditions	Text
Common Introduction (Condition 1 and 2)	LT to be sentenced in Hartford Superior court after pleading guilty to assault charges. LT, DOB. 6/21/78, is due to be sentenced in Hartford Superior Court at 2:00 p.m. on Friday, January 25, 2008 before Judge Miller. On January 16, 2008 Torres plead guilty to one count of Aggravated Battery with a Deadly Weapon. LT was arrested at 2:15 a.m. on January 3, after police responded to a 911 call regarding a violent fight in the area of Sigourney Street and Albany Avenue. While in police custody, Torres admitted to police that he stabbed his victim with a knife he was carrying after the two men became involved in a heated argument after leaving a nearby bar. The victim and several witnesses verified the account. LT was reported to be intoxicated at the time of the arrest. The victim, who sustained stab wounds to his abdomen and forearm, was transported to Hartford Hospital and is described as in serious, but stable, condition.
Cues of Association Value (Condition 2 only)	LT who manages an area auto parts store, expressed deep remorse and apologized for the pain he has caused the victim and his family. Local residents were shocked to learn of LT's arrest, saying that he had no history of causing trouble.

Notes. Subjects were randomly assigned one of the conditions. The stimuli were designed as a news release from the local police department with a picture of the criminal. The picture was randomly varied together with the name of the criminal, here abbreviated LT. These manipulations were done for purposes beyond this paper and are not analyzed in the present study.

Figure SI1. Study 2 (battery): Perceived seriousness of crime and perceived association value of the criminal as mediators of cues to association value. Structural Equation Modeling.
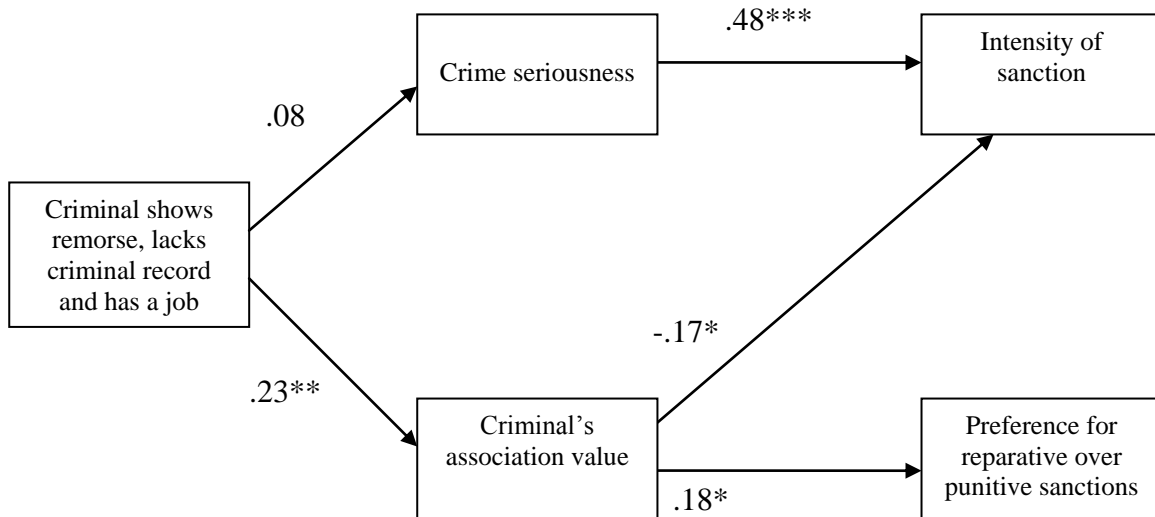
Notes. Standardized coefficients calculated from a structural equation model. Test of fitted model versus saturated model: $\chi^2 = 3.21$, p = 0.67 (two-tailed). Effect of indirect path from experimental condition through perceived association value on preference for reparative over punitive sanctions: coeff. = .04, p = .04 (one-tailed).
* p < .05, **, p < .01, *** p < .001 (one-tailed)